Title

Improved enzymatic labeling of fluorescent in situ hybridization probes applied to the visualization of retained introns in cells.

Journal

RNA, 29(8)

Authors

Xiao, Wen
Yeom, Kyu-Hyeon
Lin, Chia-Ho
et al.

Publication Date

2023-08-01

DOI

10.1261/rna.079591.123

Improved enzymatic labeling of fluorescent in situ hybridization probes applied to the visualization of retained introns in cells

WEN XIAO,[1] KYU-HYEON YEOM,[1] CHIA-HO LIN,[1] and DOUGLAS L. BLACK[1,2,3,4]

[1]Department of Microbiology, Immunology, and Molecular Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California 90095, USA

[2]Molecular Biology Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California 90095, USA

[3]Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California 90095, USA

[4]Jonsson Comprehensive Cancer Center, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California 90095, USA

ABSTRACT

Fluorescence in situ hybridization (FISH) is a widely used tool for quantifying gene expression and determining the location of RNA molecules in cells. We present an improved method for FISH probe production that yields high-purity probes with a wide range of fluorophores using standard laboratory equipment at low cost. The method modifies an earlier protocol that uses terminal deoxynucleotidyl transferase to add fluorescently labeled nucleotides to synthetic deoxyoligonucleotides. In our protocol, amino-11-ddUTP is joined to an oligonucleotide pool prior to its conjugation to a fluorescent dye, thereby generating pools of probes ready for a variety of modifications. This order of reaction steps allows for high labeling efficiencies regardless of the GC content or terminal base of the oligonucleotides. The degree of labeling (DOL) for spectrally distinct fluorophores (Quasar, ATTO, and Alexa dyes) was mostly >90%, comparable with commercial probes. The ease and low cost of production allowed the generation of probe sets targeting a wide variety of RNA molecules. Using these probes, FISH assays in C2C12 cells showed the expected subcellular localization of mRNAs and pre-mRNAs for *Polr2a* (RNA polymerase II subunit 2a) and *Gapdh*, and of the long noncoding RNAs *Malat1* and *Neat1*. Developing FISH probe sets for several transcripts containing retained introns, we found that retained introns in the *Gabbr1* and *Noc2l* transcripts are present in subnuclear foci separate from their sites of synthesis and partially coincident with nuclear speckles. This labeling protocol should have many applications in RNA biology.

Keywords: FISH; retained intron; RNA localization; alternative splicing

INTRODUCTION

RNA molecules can be localized and quantified within cells using fluorescence in situ hybridization (FISH) with complementary nucleic acid molecules (Femino et al. 1998; Raj et al. 2008; Chen et al. 2015). To reach single-molecule sensitivity these methods use strategies that target multiple dye molecules to the detected RNA molecule. One of the most straightforward methods is to tile several dozen complementary oligonucleotides along the target RNA, each carrying the fluorophore. These probes can be purchased from commercial vendors, but their cost limits the number of probe sets one can test. The limited number of available fluorophores also constrains some applications.

A protocol for the enzymatic production of fluorescently tagged DNA probes was recently described by Gaspar et al. (2017) that is inexpensive and allows efficient labeling of oligonucleotides with certain fluorophores. Some positively charged fluorophores, including ATTO488 and ATTO565, did not yield high degrees of labeling. Another limitation was that oligos with relatively high GC content yielded low labeling efficiencies presumably due to intra-

Corresponding author: dougb@microbio.ucla.edu

Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.079591.123.

and intermolecular base-pairing interactions (Gaspar et al. 2017).

We sought to apply FISH to determine the nuclear location in mammalian cells of recently characterized classes of retained introns exhibiting different fractionation behaviors (Yeom et al. 2021). Some unspliced introns are only found associated with isolated chromatin in lysed nuclei; other introns are present in transcripts of the soluble nucleoplasm; still other retained introns are present in cytoplasmic mRNAs. To characterize the subcellular and subnuclear localization of these introns, we applied smFISH using enzymatically labeled oligonucleotide probe sets. Here, we present modifications of the original protocol (Gaspar et al. 2017). With these modifications, we routinely generated probes with >90% dye labeling. Diverse commonly used fluorophores with distinct spectra including ATTO488, ATTO565, ATTO647N, Quasar570, Quasar670, and Alexa568 could all be efficiently coupled. The coupling was insensitive to the GC content and the 3′ terminal coupled nucleotide. Using this approach, we generated dozens of probe sets for roughly the cost of a single commercial probe set. We then applied these probes to locate transcripts containing different classes of retained introns relative to their sites of transcription in nuclei.

RESULTS

Oligonucleotides precoupled to amino-11-ddUTP are efficiently labeled with fluorophore regardless of GC content and terminal base

To address the variability in TdT (terminal deoxynucleotidyl transferase) labeling efficiency observed with dye-coupled ddUTP nucleotides, we changed the order of labeling steps. It was previously described that oligos carrying primary amine modifications can be labeled with succinimide-dye conjugates (Motea and Berdis 2010; Jahn et al. 2011; Winz et al. 2015), so we first added amino-11-ddUTP to the oligo 3′ ends with TdT (Fig. 1A). An oligo pool targeting the exons of *Clcn6* was designed using the Stellaris probe design tool (see Materials and Methods). To test the impact of GC content on ligation and labeling efficiency, we selected individual oligos from this pool ranging from 40% to 60% GC (the limits of the design tool) (Supplemental Fig. S1A). We also synthesized AT and GC repeat oligos (20 mer) that were 100% AT or 100% GC. After amino-11-ddUTP addition by TdT and ethanol precipitation, the oligos were incubated with NHS-ester-Quasar570 to conjugate the dye to the terminal primary amine. These labeled oligos were again precipitated and analyzed on polyacrylamide gels (Fig. 1). As shown

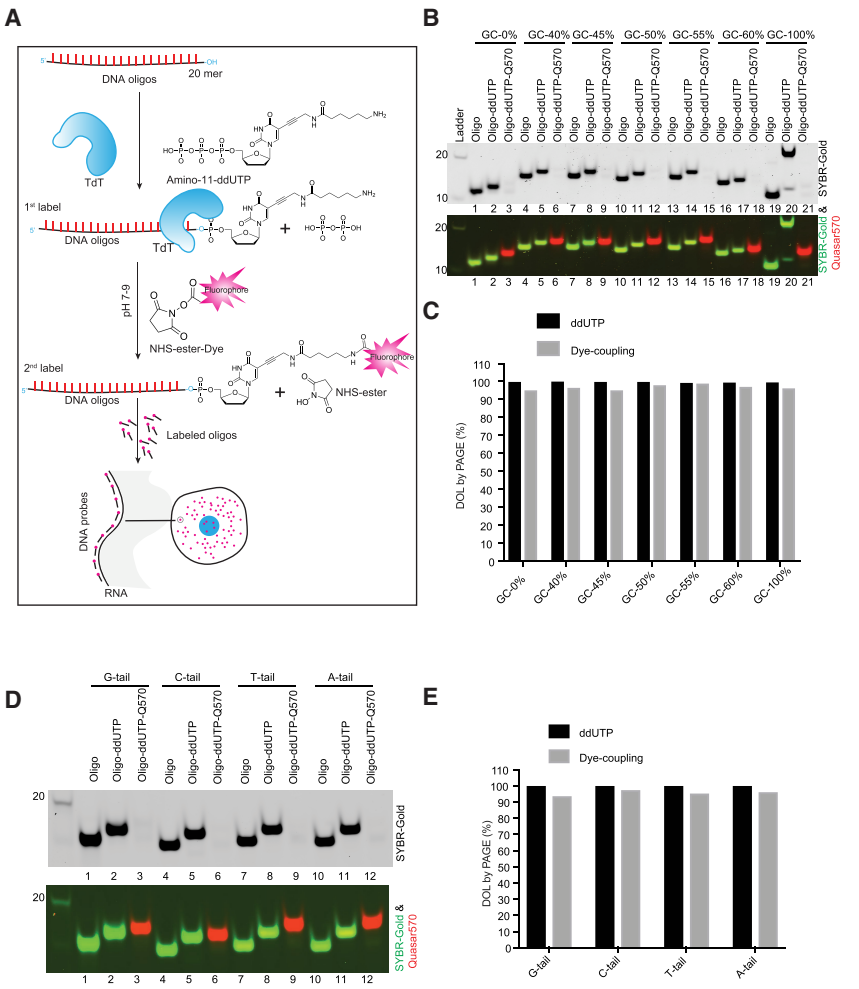

FIGURE 1. Efficient oligo coupling of amino-11-ddUTP and postlabeling with NHS-ester dyes. (*A*) Diagram of the probe labeling workflow. A set of single-stranded oligonucleotides (primers) are first coupled to amino-11-ddUTP at their 3′ end using the TdT enzyme. After ethanol precipitation, amino-11-ddUTP coupled oligos are labeled with an NHS-ester conjugated fluorescent dye. After ethanol precipitation and column purification, the oligos are used for hybridization to cells and microscopic imaging. (*B*) PAGE analysis of the degree of labeling (DOL) of ddUTP and dye for oligos of different GC content. (*Top* panel) SYBR Gold channel in gray. (*Bottom* panel) The composite of the SYBR-channel (green) and the Quasar570 channel (red). (*C*) Quantification of *B*. (*D*) PAGE analysis of first and second step labeling efficiency for oligos with four different 3′ ends. (*Top* panel) SYBR Gold channel is in gray. The *bottom* panel shows a composite of the SYBR channel (green) and the Quasar570 channel (red). (*E*) Quantification of *D*.

in Figure 1B, the TdT-modified oligos shift up in the gel, indicating the addition of one nucleotide. Across all the oligos, regardless of GC content, we observed near 100% coupling efficiency of ddUTP to the 3′ termini by the TdT enzyme (Fig. 1B,C; Supplemental Fig. S1A–C). The conjugated dye quenched staining by SYBR Gold, causing the dye-labeled band to disappear in the gel when excited at 496 nm. These species can be observed by the dye fluorescence excited at 565 and 647 nm as a band migrating slightly above the oligo coupled to ddUTP alone. The dye labeling efficiency in the second step was consistently >98% (Fig. 1C), indicating the labeling is insensitive to GC content.

TdT is known to be relatively insensitive to the 3′ terminal nucleotide it is modifying (Motea and Berdis 2010; Winz et al. 2015). To assess whether the variation of the last base influenced the dye labeling, we generated the 50% GC content oligo described above with each of the four nucleotides (A, G, C, and T) appended to the 3′ end. We found that the last base had little effect on either the ddUTP addition by TdT or the dye conjugation, with all four variants yielding a DOL above 90% (Fig. 1D,E; Supplemental Fig. S1D–F). These experiments were repeated using Quasar670 in place of Quasar570 with equivalent results, showing that two standard dyes can be used in this procedure (Fig. 1; Supplemental Fig. S1). We compared our method with that of Gaspar et al. (2017) by labeling *Neat1* probes with ATTO488. As shown in Supplemental Figure S1G,H, applying the earlier method with ddUTP precoupled to ATTO488, we obtained ∼50% and ∼70% labeling efficiency using 1× and 2× amounts of TdT enzyme, respectively. In contrast, by first coupling the amino-11-ddUTP to the oligos and then adding the succinimide-dye conjugate, we obtained >95% labeling for both ATTO488 and Quasar570 dyes (Supplemental Fig. S1I,J). These results indicate that by switching the order of addition for the ddUTP and the dye, one can obtain a high degree of probe labeling regardless of GC content and terminal base.

Pools of oligonucleotides can be labeled with equivalent efficiency to commercial FISH probe sets

A typical FISH probe set contains several dozen oligonucleotides and it is impractical to label these individually (Raj et al. 2006; Young et al. 2020). We next examined the labeling efficiency for oligo pools designed as FISH probe sets to a variety of targets. These probe sets contained 24–48 different oligos mixed in equimolar concentrations. As mixtures of sequences, these probe sets migrate as more diffuse bands than single oligos (Fig. 2). The oligo mixtures were then modified with amino-11-ddUTP using TdT and conjugated to NHS-Quasar570 or NHS-Quasar670 as above. We labeled 14 different oligo sets with Quasar570 (sets A to N) (Fig. 2A; Supplemental Table S3). We labeled 10 oligo sets with Quasar670 (sets I to VI, VI to XI) (Fig. 2D–F; Supplemental Table S4). For comparison, we purchased two Stellaris probe sets targeting the exons of *Gapdh* and the noncoding RNA-Lexis (Fig. 2A,D). We used two methods to quantify the DOL. We measured the absorbance of the dye and oligo by spectrophotometry and calculated their relative concentrations using the extinction coefficient for dye and the estimated extinction coefficient for the oligo, as done previously (Gaspar et al. 2017). As an alternative approach, we also quantified the intensities of the bands on PAGE gels and determined the DOL by comparing the SYBR Gold intensity of remaining unlabeled probes after labeling to the intensity of original oligos before labeling (Supplemental Fig. S2). Across the 25 probe sets and for both dyes, these two methods yielded very similar results, with probe sets exhibiting average DOL values of 95% and 98% in PAGE and spectrophotometric analysis, respectively (Fig. 2A,B, D,E). Using PAGE, we measured DOL values of 88% and 94.5% for the Stellaris Lexis and *Gapdh* probes, respectively, indicating that our labeling protocol is at least comparable (Fig. 2A,C,D,F).

Labeling is compatible with a range of dyes

Visualization of multiple RNA species within a single cell requires the probe set for each target to carry a dye with a different emission wavelength (Young et al. 2020). Ideally one would have a stock of amino-11-modified oligos that could be coupled to different dyes as needed. We next tested a variety of spectrally distinct NHS-modified fluorophores for their efficiency of coupling to the probe sets described above for the Quasar dyes. These fluorophores included ATTO488, ATTO565, ATTO647N, Alexa568, and Alexa647 (Supplemental Table S2). The average DOL for all the ATTO dyes was >95% according to both PAGE and spectrophotometer analysis (Fig. 3A,B; Supplemental Fig. S3A–C). Alexa568 had an average DOL >90% similar to the ATTO dyes (Fig. 3A,B, Supplemental Fig. S3D). In contrast, Alexa647 dye performed poorly with an average DOL of ∼20% in these labeling conditions (Fig. 3A,B; Supplemental Fig. S3E). Doubling the concentration of NHS-ester-Alexa647 in the labeling reaction increased the DOL to 35%. This NHS-ester-dye conjugate appears to be inherently less reactive than the others (Fig. 3A,B; Supplemental Fig. S3E).

To test the correlation of the two analytical methods, we plotted and fitted the spectroscopic and densitometric DOL values to a linear regression model (Supplemental Fig. S4A–F). As seen for the Quasar dyes, quantifying the DOL of the ATTO and Alexa dyes by either PAGE or spectrophotometry gave similar results. The spectrophotometric approach sometimes yielded DOL values >100% and were most often slightly higher than the corresponding PAGE measurements (Supplemental Fig. S4A–F). This

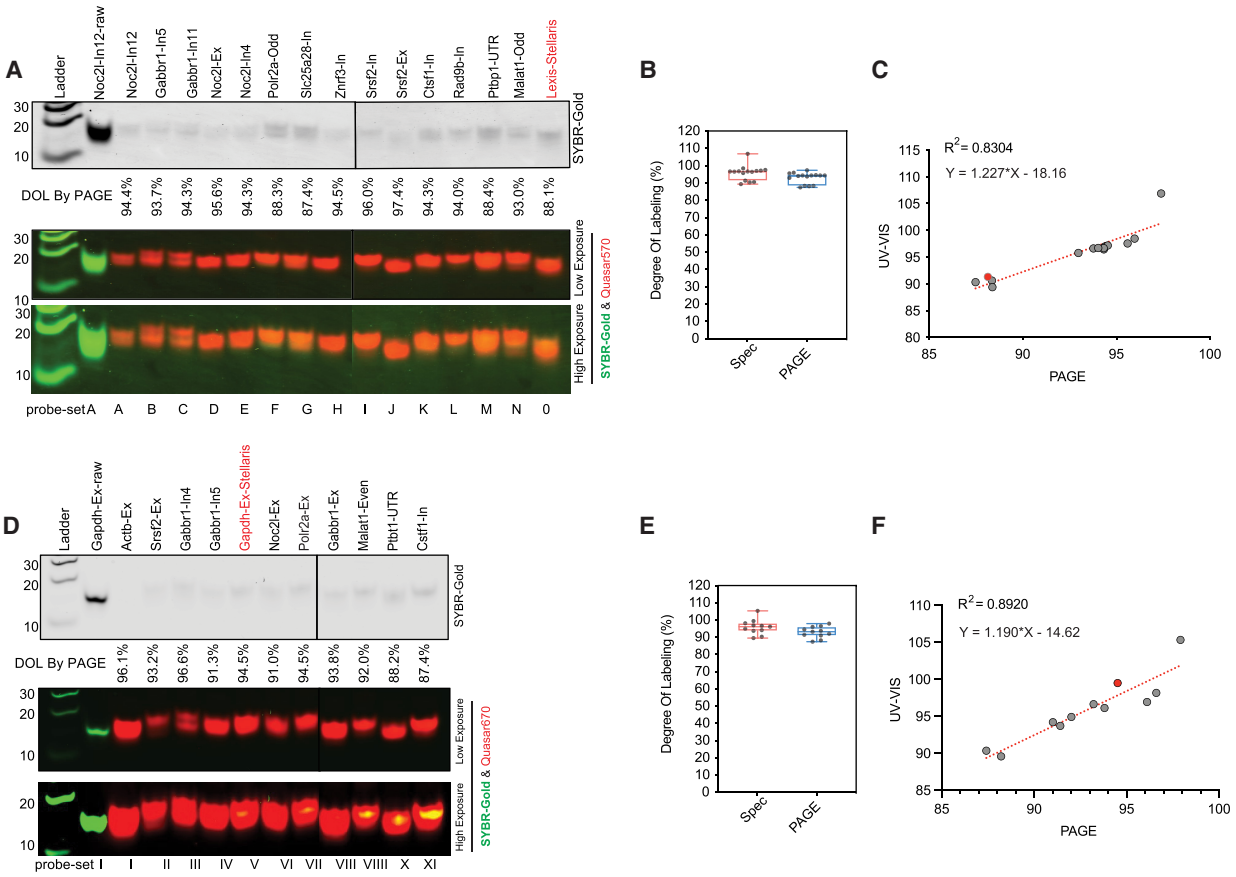

FIGURE 2. Complete FISH probe sets can be efficiently labeled in bulk with comparable efficiency to commercial probes. (*A*) PAGE analysis of Quasar570 labeling of multiple probe sets. Probe sets A to O (lanes *3–17*) contained from 15 to 48 oligonucleotides each, targeting a particular cellular transcript. Lane *2* contains the unlabeled oligos of set A. Lane *17* (set O) contains a commercial probe set from Stellaris. Band Doublets in some lanes indicate mobility differences among the mixed probes. (*B*) Box plots show the DOL of multiple probe sets determined by spectrophotometer and by PAGE analysis. Each dot represents one probe set. (*C*) The correlation of DOL measurements for Quasar570 labeling by spectrophotometer and PAGE analysis. The fitted linear equations and R^2 values are indicated. The probe set from Stellaris is highlighted in red. (*D*) PAGE analysis of Quasar670 labeling efficiency for multiple oligo sets. (Lanes *3–13*) Sets I to XI are dye-labeled oligo sets containing 15–48 oligos each. Lane *2* contains the original unlabeled oligos of set I. Lane *7* (set V) contains a commercial probe set from Stellaris. (*E*) Box plots of the Quasar670 DOL for the multiple probe sets determined by spectrophotometer (see Supplemental Fig. S2) and PAGE. (*F*) The correlation of Quasar670 DOL measurements by spectroscopy and PAGE is displayed by scatter plot. The fitted linear equations and R^2 values are indicated. The probe set from Stellaris is highlighted in red. Note that these gels were run for less time than that in Figure 1, such that the labeled and unlabeled oligos are not separated as completely.

could result from small errors in the predicted extinction coefficients for the oligo mixtures, or from changes in the dye absorption spectrum after oligo coupling. The inflation of DOL values UV–Vis measurements was also reported by other laboratories (Wang 2019). We thus believe that the DOL measurements by PAGE are more accurate. Overall, the labeling protocol functioned well with two Quasar dyes, three ATTO dyes and one Alexa dye, allowing use in multichannel FISH assays.

Performance of the probes in FISH assays

To test our probes in smFISH, we designed probe sets against a variety of cellular RNAs with previously published patterns of localization, including *Polr2a*, *Malat1*, *Neat1*, and *Gapdh*. These were designed using the Stellaris tool, synthesized, and labeled as above. The *Polr2a* transcript (RNA polymerase II subunit A) was chosen for its moderate expression and exon length (6.7 kb) that can accommodate a large number of probes (Xie et al. 2018). The 48 oligonucleotides of the *Polr2a* exon probe set were split into odd and even subsets (24 oligos each) that were coupled to different Quasar dyes (Quasar570 and Quasar670, respectively) (Fig. 4A). The odd and even probes with different fluorophores were combined prior to FISH hybridization. FISH images acquired by confocal microscopy showed odd and even spots predominantly dispersed in the cytoplasm of the C2C12 cells (Fig. 4A). Combining the odd and even channels showed visually that these spots almost entirely colocalized (Fig. 4A). Quantifying

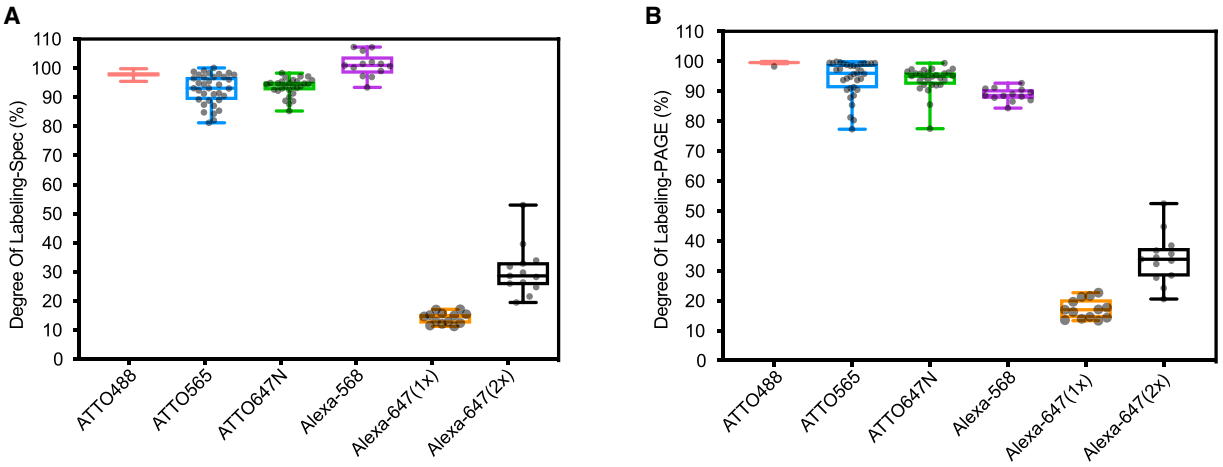

FIGURE 3. Labeling efficiency for spectrally distinct dyes. (*A*) Box plots showing the DOL of multiple probe sets with multiple ATTO and Alexa dyes determined by spectrophotometer. (*B*) Box plots of the DOL values for ATTO and Alexa dyes determined by PAGE. Each dye was used to label multiple probe sets. Each dot represents the labeling of one probe set containing up to 48 oligos (all sequences are in Supplemental Tables S5–S9). The midline of each box represents the median DOL across all the experiments, the box edges define the interquartile range and the whiskers the maximum and minimum DOL values. DOL values were determined as described in the Materials and Methods and Supplemental Figure S2.

the intensities for each pixel in ImageJ and comparing them using the JACoP plugin generated a Pearson correlation coefficient of 0.774 (Supplemental Fig. S5A). The numbers of foci observed with the odd and even probes were also quantified using the program FISH-quant (Mueller et al. 2013). The per cell spot numbers for the odd and even probes of *Polr2a* were very similar (odd: mean = 103 and even: mean = 97) (Fig. 4B), and highly correlated across many cells ($R^2 = 0.955$) (Fig. 4C). This pattern of FISH staining is nearly identical to that published previously and indicates that the probes function well to quantify mRNA within cells (Xie et al. 2018).

The Quasar dyes also performed well for localizing the long noncoding RNA *Malat1*. We created *Malat1* odd (Quasar570) and *Malat1* even (Quasar670) probes (Fig. 4D). *Malat1* is abundantly expressed and too concentrated in nuclear speckles to discern single-molecule spots for counting in FISH-quant (Fig. 4D; Fei et al. 2017; Wang et al. 2021). We measured fluorescence intensities for the odd and even channels for each pixel across the area of the nucleus in ImageJ. These per pixel intensities for the odd and even probes were highly correlated, with an average Pearson's coefficient of 89.4%. As expected, the *Malat1* probe signals in speckles were less correlated with DAPI, with the odd versus DAPI comparison yielding a Pearson's coefficient of 63%, and the even versus DAPI yielding 56.6% (Fig. 4E).

We next tested the use of dual probes to different portions of the same transcript. *Gapdh* RNA has been shown in previous studies to exhibit abundant mRNA in the cytoplasm (Rowland et al. 2019). In contrast, *Gapdh* introns are excised rapidly at the site of transcription in the nucleus (Femino et al. 1998; Rowland et al. 2019). Using ATTO565 labeled probes to target the exons and ATTO647N probes to label a long *Gapdh* intron, we confirmed this pattern (Fig. 4F). The intron probes produce two small foci per nucleus corresponding to the transcribed genes. The exon probes label these nuclear spots as well as abundant cytoplasmic mRNA (Fig. 4F). To test the Alexa568 dye, we generated two sets of probes to the lncRNA *Neat1*. *Neat1* is a component of nuclear paraspeckles and exhibits a distinctive staining pattern with several prominent nuclear foci (Ahmed et al. 2018). Probing with *Neat1* oligos labeled with both ATTO488 and Alexa568 produced the expected two to four foci per nucleus (Fig. 4G). The signals for ATTO488 and Alexa568 almost entirely overlapped, with the two dyes performing equivalently (Fig. 4H).

These experiments indicate that probes generated by this method perform well in FISH experiments using tissue culture cells.

Transcripts containing retained introns form nuclear foci distant from the site of transcription

Our goal in optimizing the FISH method was to determine the nuclear location of transcripts containing retained introns. In earlier studies, we defined different classes of retained introns based on their fractionation behavior from lysates of mouse embryonic stem cells (ESCs), neuronal progenitor cells, and cortical neurons. Several classes of introns found in nuclear polyadenylated RNA of ESCs were distinguished from classical retained introns found in cytoplasmic mRNAs (Yeom et al. 2021). Some transcripts, such as *Noc2l* pre-mRNA, contain unspliced introns in the chromatin fraction and/or the soluble nucleoplasm but are

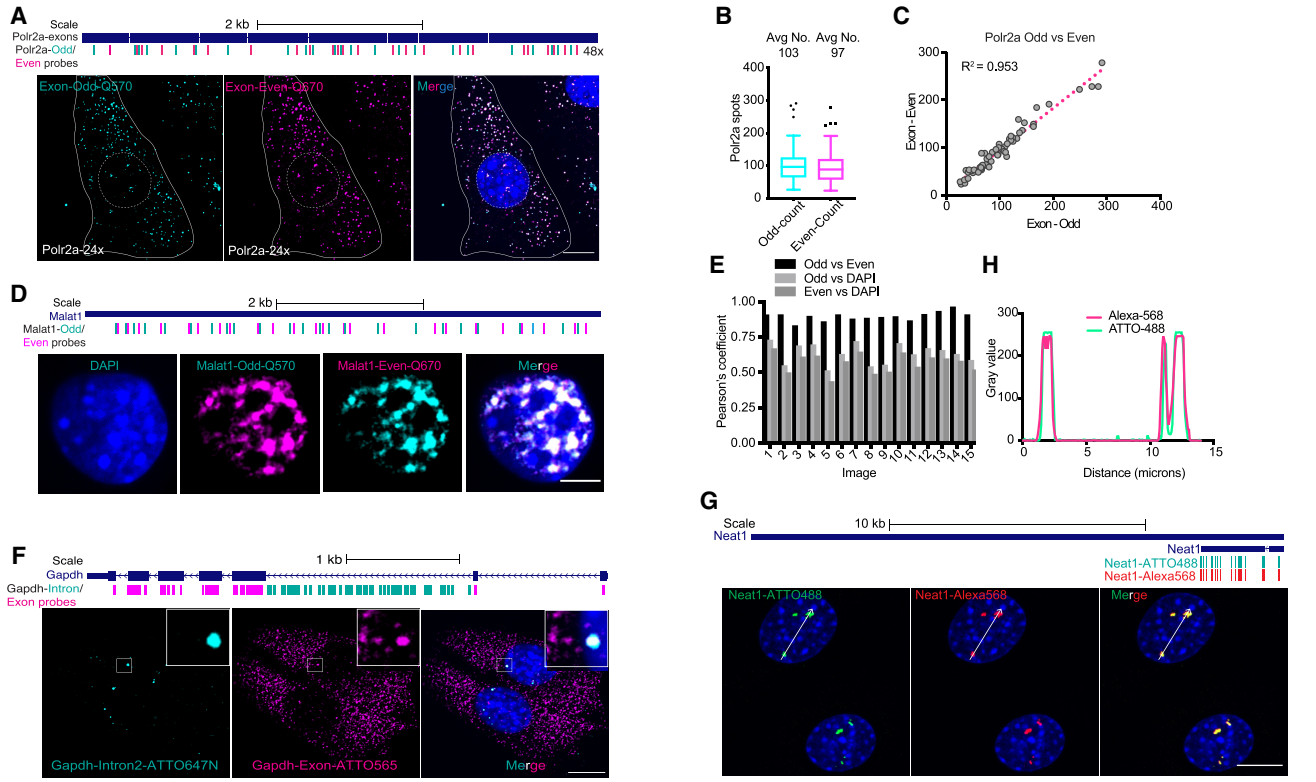

FIGURE 4. Performance of fluorescently labeled probes in FISH assays. (*A*, *top*) Diagram showing the locations of odd and even probes across the *Polr2a* coding regions (exons) in the mouse genome. (*Bottom*) Confocal microscopy images of C2C12 cells probed with the *Polr2a* odd and even FISH probes. The nuclear outline is indicated by a dashed line, and the cell boundary by a solid white line. (*B*) Box plots of the spot counts in C2C12 cells for *Polr2a* in the odd and even channels using FISH-quant software. The average number of spots per cell for the *Polr2a* odd and even signals is shown *above*. (*C*) The correlation of the odd and even spot numbers per cell are displayed in a scatter plot with the R^2 indicated. Fifty-nine cells were counted and plotted. (*D*, *top*) Diagram of the locations of the odd and even probes for *Malat1*. (*Bottom*) Confocal microscopy images of C2C12 cells probed with the *Malat1*-odd and -even probes. (*E*) Bar graph of Pearson's coefficients determined for each of 15 images. Pairwise correlations of the per pixel fluorescence intensity in two channels: Quasar570 (*Malat1*-odd) versus Quasar670 (*Malat1*-even), Quasar570 (*Malat1*-odd) versus DAPI, and Quasar670 (*Malat1*-even) versus DAPI. (*F*, *top*) Diagram of the locations of Exon (magenta) and Intron (cyan) probes for *Gapdh* arrayed along the annotated mouse gene. (*Bottom*) Confocal microscopy images of C2C12 cells hybridized with the *Gapdh* exon and intron probes. The white boxed area is enlarged at the *top right* corner. (*G*, *top*): Diagrams of the ATTO488 (green) and Alexa568 (red) probe locations arrayed along the *Neat1* short isoform in the mouse gene annotation. (*Bottom*) Confocal microscopy images of C2C12 cells hybridized with 25 *Neat1* ATTO488 probes (green), and 25 *Neat1* Alexa568 probes (red), with DAPI in blue. (*H*) Line scan profile of the ATTO488 (green) and Alexa568 (red) intensities for the cell shown in *G*. Scale bar, 10 µm.

expressed as fully spliced mRNAs in the cytoplasm (Fig. 5A). These can be classified as "detained" introns in that they are abundant in the nucleus, but are ultimately excised, and the RNAs in the cytoplasm are fully spliced (Pandya-Jones et al. 2013; Boutz et al. 2015; Mauger et al. 2016; Braun et al. 2017; Shah et al. 2018; Yeom et al. 2021). In contrast, other polyadenylated pre-mRNAs in ESCs, such as *Gabbr1*, exhibit different fractionation behavior that is similar to unpolyadenylated nascent transcripts and certain long noncoding RNAs. These transcripts are tightly associated with the pelleted chromatin, but found only in limited amounts in the soluble nucleoplasm and cytoplasm (Fig. 5E; Yeom et al. 2021). Introns in these RNAs remain unexcised and the transcripts do not produce mature mRNA. Speculating that these RNAs might be anchored at their sites of transcription, we applied FISH to examine to the location of nuclear polyadenylated RNAs containing highly retained introns.

To examine spliced and unspliced *Noc2l* transcripts in ESCs, we generated separate FISH probe sets for the *Noc2l* exons, and for introns 4 and 12 (Fig. 5A,B). Intron 4 is highly spliced in all compartments, whereas intron 12 is largely unspliced in the chromatin fraction, present at lower levels in the soluble nucleoplasm, and fully excised in the cytoplasm. Mouse ES cells probed for the *Noc2l* exons showed abundant spots in the cytoplasm as expected, as well as multiple foci in the nucleus. Cells probed for the rapidly spliced intron 4 showed only zero, one, or two nuclear foci that also stained with exon probes and are presumably the sites of transcription (Fig. 5A,B; Supplemental Fig. S6B). Intron 4 staining was not observed in the cytoplasm in agreement with the sequencing data (Fig. 5A;

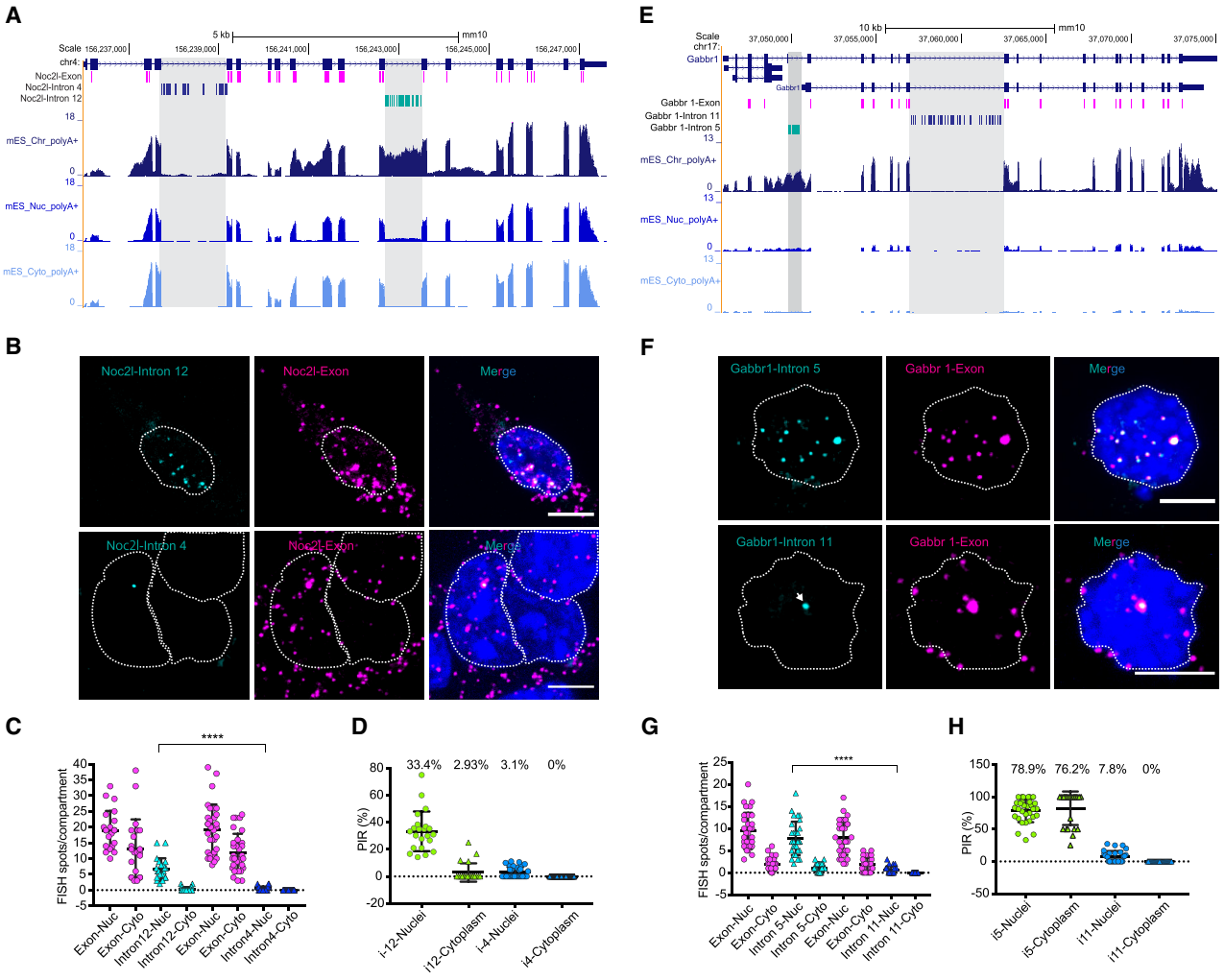

FIGURE 5. Localization of retained introns in mouse ES cell nuclei. (*A*) UCSC genome browser tracks show *Noc2l* poly(A)$^+$ RNA from the chromatin, nucleoplasmic, and cytoplasmic fractions. (*Top*) The genomic structure of the *Noc2l* gene is diagrammed with exon and intron probe locations in magenta, blue, and green. (*Bottom*) RNA-seq reads from the three fractions are displayed. Introns probed by FISH are indicated in gray. (*B*) FISH images of mES cells probed for *Noc2l* intron 12 (ATTO647N, retained), intron 4 (ATTO647N, unretained), and exons (ATTO565). Nuclear borders are outlined in white dashed lines. (*C*) Quantification of spot counts for the *Noc2l* exon and intron probes in the nuclear and cytoplasmic compartments. (*D*) Percentage intron retention (PIR) for *Noc2l* introns 12 and 4 in nuclear and cytoplasmic compartments as determined by the ratio of intron/exon costaining spots to the total exon spots in mES cells. (*E*) UCSC genome browser tracks show *Gabbr1* poly(A)$^+$ RNA from the chromatin, nucleoplasmic, and cytoplasmic fractions. (*Top*) The genomic structure of the *Gabbr1* gene is diagrammed with exon and intron probe locations in green, blue, and magenta. (*Bottom*) RNA-seq reads from the three fractions are displayed. Introns probed by FISH are indicated in gray. (*F*) FISH images of mES cells probed for *Gabbr1* introns 5 (ATTO647N, retained) and 11 (ATTO647N, unretained) in green and for the *Gabbr1* exons (ATTO565) in magenta. Nuclear borders are outlined in white dashed lines. (*G*) Quantification of spot counts for the *Gabbr1* exon and intron probes in the nuclear and cytoplasmic compartments. (*H*) PIR for *Gabbr1* introns 5 and 11 in nuclear and cytoplasmic compartments as determined by the ratio of intron/exon costained spots to total exon spots in mES cells. Scale bar, 5 µm.

Supplemental Fig. S6B). In contrast to intron 4, the probes for the highly retained intron 12 colocalized with nearly all the nuclear exon foci whether they were at the sites of transcription or not. These intron 12 foci were absent from the cytoplasm (Fig. 5C,D; Supplemental Fig. S6A). Some nuclear *Noc2l* exon puncta did not stain for intron 12 and are likely to be spliced mRNAs. Intron puncta that did not costain for the exons were much fewer in number and may be excised intron 12 lariats that are not yet degraded. Overall, *Noc2l* transcripts containing a retained intron 12 do not remain at the site of transcription, and are spliced after release from the gene locus but before export to the cytoplasm. This is in agreement with prior studies of introns that are spliced post-transcriptionally (Girard et al. 2012; Boutz et al. 2015; Coté et al. 2020; Barutcu et al. 2022).

We next examined the localization of the *Gabbr1* transcripts that exhibit a different pattern of splicing and fractionation. In mouse ESCs, *Gabbr1* transcripts are almost entirely unspliced in a region of complex processing encompassing intron 5 (Fig. 5E). Other *Gabbr1* introns

including intron 11 appear spliced even in the chromatin fraction. We generated *Gabbr1* probes for the exons, for intron 11, and for the region of intron 5 present in the major isoform expressed in mESCs (Fig. 5E). All three probe sets yielded very little staining in the cytoplasm in agreement with the fractionated RNA-seq data indicating that these RNAs remain in the nucleus. The limited spots observed in the cytoplasm usually stained for both the exons and intron 5. Exon foci for *Gabbr1* were more abundant in the nucleus, with most cells exhibiting one large spot per nucleus and five to 15 smaller spots. Nearly all these exon foci costained for intron 5 (Fig. 5F; Supplemental Fig. S6C). In contrast, the probes for the rapidly excised intron 11 yielded only one spot per nucleus, which is presumably the site of transcription (Fig. 5F; Supplemental Fig. S6D). The intron 11 staining was seen in the center of the largest spot costained by the exon and intron 5 probes. These large foci may indicate an abundance of exon and intron 5 RNA concentrated in the region of the gene locus, although there were multiple smaller foci away from this site that also costained for the exons and intron 5.

The spot numbers for introns 5 and 11 of *Gabbr1* and for introns 4 and 12 of *Noc2l* in the nucleus were quantified (Fig. 5C,G). The rapidly excised introns 12 and 4 averaged less than two foci per nucleus, consistent with them marking the gene loci. In contrast, *Noc2l* intron 11 and *Gabbr1* intron 5, which exhibit much lower levels of excision, generated many more nuclear foci, whose numbers approached the numbers of exon foci. Computing the PIR from the ratio of costained intron/exon foci to the total exon foci yielded values in agreement with those calculated from sequencing data (Fig. 5D,H; Supplemental Fig. S6E). Although the foci at the transcriptional loci were larger, *Gabbr1* RNA localization was similar to *Noc2l*, with multiple foci scattered through the nucleus away from the site of transcription (Fig. 5G,H). These data indicate that despite differences in their fractionation behavior and eventual splicing, RNAs containing retained introns are being released from both of these gene loci. In the case of *Noc2l*, these RNAs are eventually spliced and exported to the cytoplasm, whereas for *Gabbr1* the RNAs remain unspliced and unexported.

We also created exon and intron probe sets for five other genes containing detained introns in ESCs: *Cstf1*, *Neil3*, *Rad9b*, *Slc25a28*, and *Znrf3* (Supplemental Fig. S5). As seen in the sequencing tracks (Supplemental Fig. S5C,E, G,I,K), these genes express polyadenylated RNA containing an intron in the chromatin and/or nucleoplasmic fractions but are fully spliced in the cytoplasm. *Cstf1* and *Slc25a28*, which are the most highly expressed, exhibited abundant exon foci including many in the nucleus. Some of these nuclear foci costained for the retained intron as expected. *Neil3*, *Rad9b*, and *Znrf3* were expressed at lower levels in these cells. The nuclear exon foci for these genes were fewer in number but usually included one or two larger puncta that were presumably the sites of transcription. These larger puncta also costained with the intron probes for *Neil3* and *Rad9b*. The intron probe set for *Znrf3* did not perform well and yielded only nonspecific background. Overall, these genes exhibited nuclear RNA fractionation and nuclear intron/exon staining similar to *Noc2l*.

Highly retained introns partially coincident with nuclear speckles

Nuclear speckles are regions within the nucleoplasm enriched for splicing factors, such as SRSF1 and SRSF2 (SC35), and for the long noncoding RNA *Malat1*. Active gene loci are often observed at the periphery of speckles, and post-transcriptional splicing is observed within them (Dias et al. 2010; Girard et al. 2012; Galganski et al. 2017). To confirm the speckled staining pattern of the *Malat1* FISH probes, we combined *Malat1* FISH with immunofluorescence using SRSF2 (SC35) and SRSF1 antibodies in C2C12 cells (Supplemental Fig. S7). *Malat1* staining strongly overlapped with SRSF2 (SC35) (Supplemental Fig. S7A,B). Cross-sectional profiles through the nucleus presented largely overlapping peaks for intensity for the two signals (Supplemental Fig. S7B). SRSF1 showed partially coincident staining with *Malat1* with many intense puncta costaining for both protein and RNA, but also exhibited regions of protein staining that did not overlap with *Malat1* localization (Supplemental Fig. S7C,D).

Using *Malat1* as a speckle marker, we examined whether the retained intron transcripts are within speckles. For *Noc2l* and *Gabbr1*, we used separate labels for the retained intron and exon probes with a third fluorophore on the *Malat1* probes in multiplexed FISH assays. Identifying foci of unspliced introns defined by colocalized exon and intron sequences, we observed approximately five colocalized exon–intron spots for *Noc2l* per nucleus and approximately 10 such spots for *Gabbr1* (Fig. 6A,B, D,E). We then measured the overlap of these retained intron foci with *Malat1* staining in speckles. We found that the majority of these retained intron transcripts colocalized with *Malat1* (Fig. 6A,B,D,E). On average, 70% of the *Noc2l* intron 12 and 76% of the *Gabbr1* intron 5 spots overlapped with *Malat1* staining (Fig. 6C,F). For both introns, there was a subset of transcripts that did not appear to be in speckles. Simultaneous probing for *Noc2l* intron 12 and *Gabbr1* intron 5 did not show overlap (Fig. 6G), indicating that speckles are not homogeneous but instead contain different RNAs in different regions.

DISCUSSION

Labeling oligonucleotide probe sets

We present an improved method for the rapid, efficient and inexpensive production of fluorescent oligonucleotides.

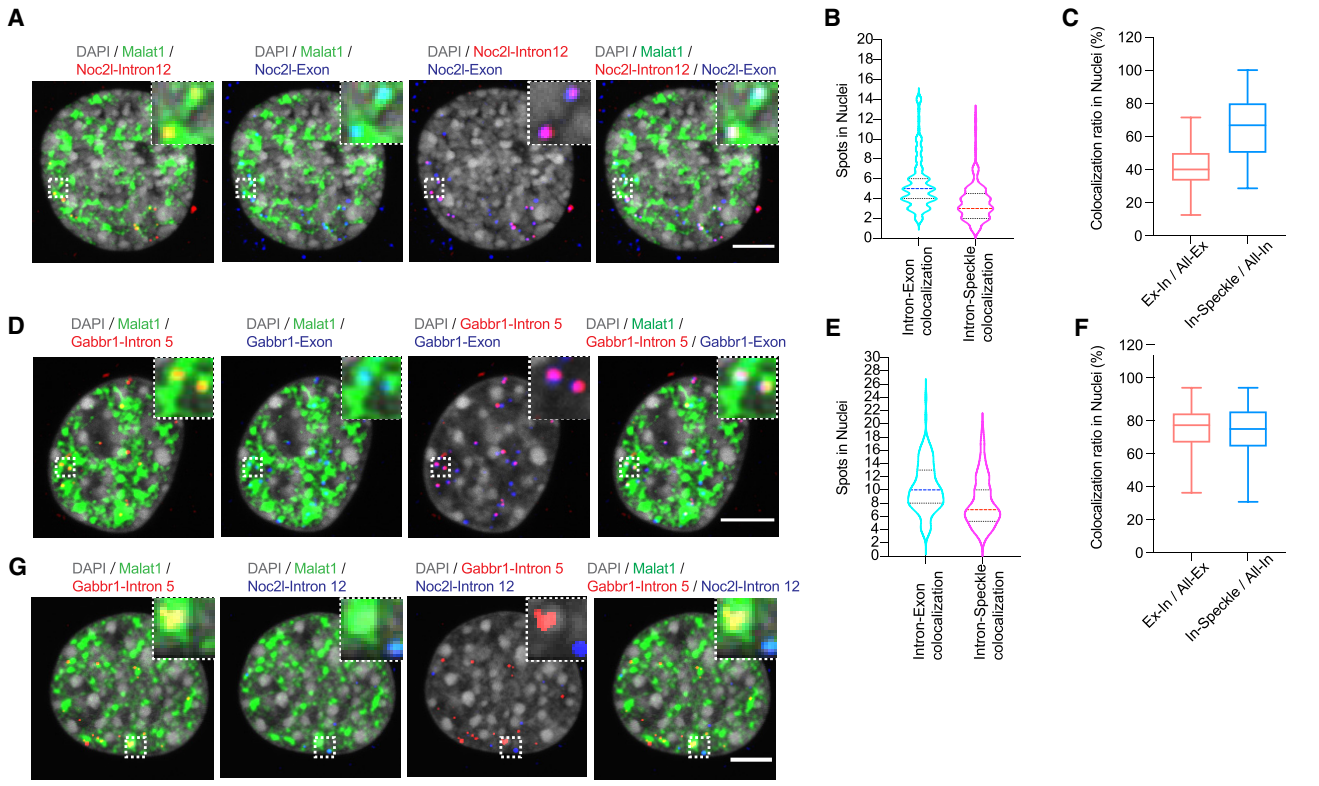

FIGURE 6. Retained intron transcripts of *Gabbr1* and *Noc2l* are dispersed in the nucleus and partially colocalize with speckles. (*A*) FISH images of C2C12 cells simultaneously probed for *Malat1* (ATTO488, green), *Noc2l* intron 12 (ATTO647N, red), and *Noc2l*-exon (ATTO565, blue). DAPI stain is shown in gray. (*B*) Violin plots showing numbers and distribution of colocalized *Noc2l* exon and intron spots per cell, and numbers of intron spots colocalized with speckle signal per cell, across 101 cells. (*C*) Box plots showing the ratios of the colocalized *Noc2l* intron–exon spots over the total exon spots in nucleus (In-Ex/All-Ex) per cell, and the intron spots with overlapping speckle (*Malat1*) signal over the total intron spots in nucleus (In-Speckle/All-In) per cell, across 101 cells. (*D*) FISH images of C2C12 cells simultaneously probed for *Malat1* (ATTO488, green), *Gabbr1* intron 5 (ATTO647N, red), and *Gabbr1* exons (ATTO565, blue). DAPI stain is shown in gray. (*E*) Violin plots showing numbers and distribution of colocalized *Gabbr1* exon and intron spots per cell, and numbers of intron spots colocalized with speckle signal per cell, across 104 cells. (*F*) Box plots showing the ratio of the colocalized *Gabbr1* intron–exon spots over the total exon spots in nucleus per cell, and the intron spots with overlapping speckle (*Malat1*) signal over the total intron spots in nucleus per cell, across 104 cells. (*G*) FISH images of C2C12 cells simultaneously probed for *Malat1* (ATTO488, green), *Gabbr1* intron 5 (ATTO565, red), and *Noc2l* intron 12 (ATTO647N, blue), with DAPI stain in gray.

This protocol is derived from an earlier one using terminal deoxynucleotidyl transferase to add functional moieties to the 3′ end of a standard DNA oligo (Gaspar et al. 2017). We found that switching the order of addition and making several smaller adjustments, significantly improved the labeling efficiency and allows the use of a broader range of fluorophores and oligonucleotide sequences.

In the new procedure, amino-11-ddUTP is first coupled to an oligonucleotide pool using TdT. This step is very efficient and allows the creation of a stock probe set that can be subsequently labeled with the fluorophore of choice. The oligo pools are then treated with a succinimide-dye conjugate to add the fluorophore. For nearly all dyes tested, this labeling is also very efficient, and not affected by either the GC content or 3′ terminal nucleotide of the probe. Comparing Quasar dye-coupled laboratory-made probes with commercial Quasar probes, we found that they had comparable labeling efficiencies and performed equivalently. Besides the Quasar dyes, we successfully labeled probes with commonly used ATTO dyes (ATTO488, ATTO565, and ATTO647N) and an Alexa dye (Alexa568) enabling many combinations for multiplexed imaging (Supplemental Figs. S3, S4; Supplemental Tables S2, S5–S9). One dye, Alexa647, did not efficiently couple with the amino group and was not usable. It will be interesting to test additional dyes and other modifications such as biotin for coupling to our oligonucleotide probe sets.

In the course of these experiments, we generated more than 100 probe sets containing 24–48 oligos each and labeled with six different dyes. With these materials, the labeling of one probe set costs about $15, and can be used in approximately 1000 FISH assays (Supplemental Table S1). This is compared with the approximate $800 cost of a single commercial probe set of 48 oligonucleotides labeled with a Quasar dye.

The use of tiled oligonucleotides to target multiple fluorophores to an RNA molecule by in situ hybridization has been a standard strategy for in vivo imaging experiments.

Several other cost-effective FISH protocols have been developed. The single-molecule inexpensive FISH (smiFISH) uses unlabeled primary probes against the RNA target along with dual fluorescent probes binding to the primary probe for signal detection (Tsanov et al. 2016; Calvo et al. 2021). Another method, termed single-molecule hybridization chain reaction (smHCR), uses unlabeled primary probes to the target RNA and then chain reaction probes that can be labeled and amplified (Shah et al. 2016). Another enzyme-based FISH method uses phagemid to produce ssDNA probes. The purified and DNase I-digested ssDNAs (normally 20–60 nt) are then used for enzyme-based probe synthesis (Lanctôt 2018). These methods each have their own potential advantages. It will be useful to compare them with standard oligonucleotide FISH now that their costs are more equivalent.

Localizing retained introns

Most introns are excised during transcription of the nascent RNA and localize to the gene locus in FISH analyses (Shah et al. 2018). Other introns are excised after transcription is complete. Some of these unspliced transcripts are released from the DNA template and can be observed in the nucleoplasm at a distance from the gene and sometimes in the cytoplasm (Vargas et al. 2011; Girard et al. 2012; Coulon et al. 2014; Coté et al. 2020). In earlier studies, we and others have characterized retained introns and their regulation (Frankiw et al. 2019; Monteuuis et al. 2019; Grabski et al. 2021; Yeom et al. 2021). Classical retained introns are found in transcripts that have been exported to the cytoplasm as mature mRNAs. Other introns exhibit delayed splicing and are abundant in nuclear RNA but are excised prior to the export of the fully processed RNA. These "detained" introns have been observed in inflammatory response transcripts of mouse macrophages, in transcripts affecting cell cycle regulation and growth control in human cell lines, in cultured postmitotic neurons and in mouse brain (Bhatt et al. 2012; Pandya-Jones et al. 2013; Wong et al. 2013; Boutz et al. 2015; Mauger et al. 2016; Frankiw et al. 2019). These introns are thought to act as brakes that slow the expression of the transcripts where they are found. In some cases, this may create a pool of RNAs that can be rapidly matured into mRNA in response to a signal. In different systems, reduced excision rates have been ascribed to reduced activity of the arginine methyltransferase PRMT5, and to the presence or absence of particular RNA binding proteins (Hayashi et al. 2014; Braun et al. 2017; Frankiw et al. 2019). Other studies found that one class of retained introns was derived from genes associated with the nuclear periphery and lamina, whereas another class including most detained introns was expressed from more centrally located genes whose transcripts colocalized to nuclear speckles (Barutcu et al. 2022; Tammer et al. 2022).

We recently characterized retained introns based on their fractionation into cytoplasmic and two subnuclear compartments (Yeom et al. 2021). Within the nuclei, we identified both nascent and polyadenylated RNAs that were tightly associated with the high molecular weight pellet fraction containing chromatin and some of the nuclear speckle material. This fraction was distinguished from a soluble nuclear fraction that contains mature mRNAs in transit to the cytoplasm, as well as some polyadenylated but incompletely spliced RNAs. Nuclear introns included the described detained introns that were found both in the chromatin pellet and the soluble nucleoplasm, but were absent from mature RNA in the cytoplasm.

We also identified an unusual class of nuclear polyadenylated transcripts that contain introns but are absent from the cytoplasm as either spliced or unspliced RNA. These transcripts were tightly associated with the chromatin pellet, showing fractionation behavior similar to Xist and other chromatin-bound lncRNAs (McHugh et al. 2015; Mishra and Kanduri 2019; Pandya-Jones et al. 2020; Yeom et al. 2021). For these RNAs, the failure to splice and possibly their tight association with chromatin prevents the expression of the encoded protein. The most notable example is *Gabbr1*, which is transcribed in ESCs but fails to splice intron 5; *Gabbr1* transcripts remain tightly associated with the chromatin pellet and, despite substantial transcription, fail to produce *Gabbr1* protein. As cells differentiate into neurons, these *Gabbr1* gene RNAs become fully processed to produce cytoplasmic mRNA and *Gabbr1* protein (Yeom et al. 2021). The unusual fractionation behavior of the *Gabbr1* transcripts in ESCs and its similarly to Xist, led to the hypothesis that these transcripts might condense at their site of transcription rather than being released into the nuclear speckle as seen with RNAs destined to be fully spliced, such as detained intron transcripts.

To examine the location of *Gabbr1* transcripts we developed FISH probe sets for the retained *Gabbr1* intron 5, for the *Gabbr1* exons, and for *Gabbr1* intron 11. Intron 11 is excised cotranscriptionally and unlike intron 5 is only seen in the nascent unpolyadenylated *Gabbr1* RNA. The localization patterns of the *Gabbr1* RNAs were compared with those of *Noc2l*. *Noc2l* intron 12 is also highly retained in the nuclear RNA but is eventually excised and, unlike *Gabbr1*, the mature *Noc2l* mRNA is exported and translated. The localization of *Noc2l* intron 12 was compared with the *Noc2l* exons and *Noc2l* intron 4, which is cotranscriptionally excised. Probing for the cotranscriptionally excised introns identified the sites of the active gene loci, which were also labeled with exon and retained intron probes for both genes. In contrast to the cotranscriptional introns, the retained introns for both the *Gabbr1* and *Noc2l* genes, were also seen in foci distant from the sites of transcription that colabeled with the exon probes. Thus, both of these retained intron transcripts are being released from the

DNA template but are staining at different locations. The *Noc2l* transcripts will eventually be spliced but the *Gabbr1* transcripts are presumably destined for degradation. A portion of the fluorescent signal for each of these transcripts colocalized with the speckle marker *Malat1*, consistent with previous results showing retained introns in speckles, while a portion of each did not. Overall, we did not observe differences in the localization of the *Gabbr1* and *Noc2l* transcripts that might correlate with their different fractionation behaviors. Interestingly, the foci at the transcription site for *Gabbr1* are noticeably larger than seen for *Noc2l* or other transcripts, possibly an indication that it forms a different molecular condensate.

MATERIALS AND METHODS

FISH probe production and purification

Oligo design and purchase

Arrays of 20-mer DNA oligos complementary to RNA targets were designed using the Stellaris Probe Designer, version 4.2 (https://www.biosearchtech.com/stellaris-designer). Selection criteria included 40%–60% GC content and similar melting temperatures. The spacing length was at least 2 nt between oligos. Oligos were purchased from IDT as desalted oligos in 96-well plates at 200 µM concentration in nuclease-free water. Oligos were thawed and spun to remove debris before labeling.

Labeling oligo pools with amino-11-ddUTP

Probe set pools were assembled by mixing 10 µL from each oligo well. Enzymatic labeling was performed in a 40-µL mixture containing 28 µL of the oligo pool (200 µM in H_2O), 1 µL of amino-11-ddUTP (10 mM in H_2O; Lumiprobe), 8 µL of 5× TdT buffer (Thermo Scientific), and 3 µL of 20 U/µL TdT enzyme (Thermo Scientific). Labeling reactions were incubated overnight at 37°C in a PCR thermocycler with a hot lid. Reactions were stopped by incubation for 10 min at 70°C. The oligos were then ethanol-precipitated by the addition of 4 µL of 3 M sodium acetate (pH 5.2; AMRESCO E521-100 mL) and 110 µL of 100% ethanol and cooled for at least 1 h to −80°C. The samples were then centrifuged at 14,000 rpm for 15 min to pellet the oligos. The oligo pellets were washed twice with cold 75% ethanol, air-dried, and dissolved in 15 µL of $NaHCO_3$ (0.1 M in H_2O at pH 8.3).

Coupling NHS-ester dyes to the oligos

The dye labeling reactions included 15 µL of amino-modified oligos in $NaHCO_3$ from above and 0.75 µL of succinimide-dye conjugate (20 mM, dissolved in DMSO). The mixture was incubated for 2 h at room temperature (RT) in the dark. Oligos were ethanol-precipitated by adding the following to each 15.75 µL of reaction mixture: 34.25 µL of nuclease-free water, 5 µL of 3 M sodium acetate, and 137.5 µL of 100% ethanol; cooling for 1 h to −80°C; and centrifuging at 14,000 rpm for 15 min. The now colored pellets were washed three times with prechilled 75% ethanol, air-dried, and dissolved in 50 µL of nuclease-free water or Tris-EDTA buffer (10 mM Tris-HCl, 1 mM EDTA).

Column purification

To eliminate free fluorescent dyes, the labeled oligos went through column purification. We added 5 µL of 3 M sodium acetate (pH 5.2) and 137.5 µL of 100% ethanol to 50 µL of dissolved dye-labeled oligos and cooled the mixture for at least 1 h at −80°C. The cold mixture was quickly transferred to a PCR clean-up column (Promega A9282) and centrifuged at 14,000*g* for 30 sec. This step should be done very quickly, as warming of the tube may enhance the solubility of the oligos and reduce their binding to column membrane. Columns carrying the membrane-retained probes were washed two times with prechilled 80% ethanol and then spun once to remove extra ethanol. An amount of 50 µL of nuclease-free water or Tris-HCl buffer was added to elute the labeled probes. The probe concentrations were measured by Nanodrop, and used for FISH experiments.

NHS-ester dyes dissolved in DMSO at a concentration of 10 mM were stored in the dark at −20°C. Dye-labeled oligos were dissolved in TE buffer and stored in small aliquots at −20°C. Both reagents were stable in dark storage for up to a year before decay of the absorbance properties became apparent.

Quantification of degree of labeling

DOL of the labeled oligonucleotides was analyzed by UV–Vis spectroscopy using a Nanodrop spectrophotometer and by PAGE.

The DOL of labeled oligos by spectroscopy was estimated using the following formula:

$$\mathrm{DOL(UV-Vis)} = \frac{\mathrm{OD_{dye}}}{\varepsilon_{\mathrm{dye}}} \times \frac{(\varepsilon_{\mathrm{oligo}} + 9000)}{(\mathrm{OD_{oligo}} - \mathrm{cf_{260}} \times \mathrm{OD_{dye}})}.$$

Absorbance at 260 nm (OD_{oligo}) and at the dye maximal absorbance (OD_{dye}) was measured by Nanodrop, and the above formula was used to estimate the properties of the labeled oligonucleotides. ε_{dye} is the extinction coefficient of dye and the cf_{260} represents the relative absorbance of dye at 260 nm. The ε_{oligo} is calculated by taking an average of the extinction coefficients of the oligos in the mixture and adding 9000 mol/cm corresponding to the incorporated ddUTP. The DOL of labeled oligos by PAGE was determined using the following formula:

$$\mathrm{DOL(PAGE)} = 1 - \frac{U}{\mathrm{SO}},$$

where U stands for the band intensity by SYBR staining of the remaining unlabeled oligos after dye labeling, and SO stands for the band intensity by SYBR staining of the starting oligos before labeling.

Experimental protocol for FISH

C2C12 cells or mouse ES cells were cultured on gelatin-coated coverslips. The cells were washed once with ice-cold PBSM (1× PBS, 5 mM $MgCl_2$), followed by fixation with 4% paraformaldehyde in PBSM for 10 min at RT. After a 5-min wash with ice-cold PBSM, the cells were immersed in 70% ethanol for permeabilization overnight. The ethanol was then removed and 10% formamide in 2× SSC buffer (300 mM sodium chloride, 30 mM sodium

citrate) was added for 2 h to rehydrate and predenature the RNAs. The probes were suspended in hybridization buffer (Biosearch SMF-HB1-10; 10% formamide added freshly) to a final concentration of 0.5 ng/µL. The coverslips were removed from the dish, dipped on a Kim wipe to remove extra buffer, and then flipped onto droplets of hybridization buffer on parafilm. Hybridization was carried out in humidified chambers for at least 16 h at 37° C. The samples were then washed twice with wash buffer A (Biosearch SMF-WA1-60) for 30 min at 37°C. DAPI (Invitrogen D1306) was added in the second wash to a final concentration of 0.5 µg/mL. The cells were then washed once with wash buffer B (Biosearch SMF-WB1-20) for 5 min. The coverslips were mounted in ProLong Gold antifade media (Invitrogen P36930) overnight before microscopy.

Image analysis

Multicolor *z*-stack images acquired on a Leica confocal SP8 microscope were split to individual channels and imported into FISH-quant (Mueller et al. 2013). Nuclei and cytoplasm were segmented by drawing outlines manually using DAPI and fluorescent dye signals. Filtering and local focus projection parameters were optimized per data set. For *Polr2a* transcript detection, we followed the workflow of FISH-quant for spot detection (Mueller et al. 2013). Images were filtered, predetection was performed, spots were filtered, and fits were further thresholded to exclude outliers. All analysis was performed via the FISH-quant batch processing toolbox. The JACop plugin in ImageJ was used for colocalization analysis of the *Polr2a* odd and even FISH images. For retained intron FISH analysis, the multicolored images were first split into individual channels, then the *z*-stacked single channel was max-projected through ImageJ. The nuclear and cell outlines were manually drawn using DAPI and fluorophore signals. The nuclear and cytoplasmic spots were counted manually based on those outlines. Line scan profiles were generated using the "Plot Profile" module in ImageJ.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank Imre Gaspar, Anne Ephrussi, Tyler Matheny, Roy Parker, Yolanda Markaki, Kathrin Plath, Tamer Sallam, and Fangyuan Ding for advice and materials for oligo labeling and FISH; Mohammad Nazim for his critical reading of the manuscript and helpful discussion. This work was supported by National Institutes of Health grants R35 GM136426, R01 CA220238, and R01 GM114463 to D.L.B.; a grant from the W.M. Keck Foundation to D.L.B.; and grants from the Broad Stem Cell Research Center and COVID-19 Research Grant Program at University of California, Los Angeles to D.L.B.

Author contributions: W.X. and D.L.B. designed the research; W.X. and K.-H.Y. performed the experiments; W.X., C.-H.L., and D.L.B. analyzed the data; and W.X. and D.L.B. wrote the paper.

Received January 10, 2023; accepted March 31, 2023.

REFERENCES

Ahmed ASI, Dong K, Liu J, Wen T, Yu L, Xu F, Kang X, Osman I, Hu G, Bunting KM, et al. 2018. Long noncoding RNA *NEAT1* (nuclear paraspeckle assembly transcript 1) is critical for phenotypic switching of vascular smooth muscle cells. *Proc Natl Acad Sci* **115:** E8660–E8667. doi:10.1073/pnas.1803725115

Barutcu AR, Wu M, Braunschweig U, Dyakov BJA, Luo Z, Turner KM, Durbic T, Lin Z-Y, Weatheritt RJ, Maass PG, et al. 2022. Systematic mapping of nuclear domain-associated transcripts reveals speckles and lamina as hubs of functionally distinct retained introns. *Mol Cell* **82:** 1035–1052.e9. doi:10.1016/j.molcel.2021.12.010

Bhatt DM, Pandya-Jones A, Tong A-J, Barozzi I, Lissner MM, Natoli G, Black DL, Smale ST. 2012. Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell* **150:** 279–290. doi:10.1016/j.cell.2012.05.043

Boutz PL, Bhutkar A, Sharp PA. 2015. Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev* **29:** 63–80. doi:10.1101/gad.247361.114

Braun CJ, Stanciu M, Boutz PL, Patterson JC, Calligaris D, Higuchi F, Neupane R, Fenoglio S, Cahill DP, Wakimoto H, et al. 2017. Coordinated splicing of regulatory detained introns within oncogenic transcripts creates an exploitable vulnerability in malignant glioma. *Cancer Cell* **32:** 411–426.e11. doi:10.1016/j.ccell.2017.08.018

Calvo L, Ronshaugen M, Pettini T. 2021. smiFISH and embryo segmentation for single-cell multi-gene RNA quantification in arthropods. *Commun Biol* **4:** 352. doi:10.1038/s42003-021-01803-0

Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X. 2015. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348:** aaa6090. doi:10.1126/science.aaa6090

Coté A, Coté C, Bayatpour S, Drexler HL, Alexander KA, Chen F, Wassie AT, Boyden ES, Berger S, Churchman LS, et al. 2020. pre-mRNA spatial distributions suggest that splicing can occur post-transcriptionally. *Cell Biol* doi:10.1101/2020.04.06.028092

Coulon A, Ferguson ML, de Turris V, Palangat M, Chow CC, Larson DR. 2014. Kinetic competition during the transcription cycle results in stochastic RNA processing. *Elife* **3:** e03939. doi:10.7554/eLife.03939

Dias AP, Dufu K, Lei H, Reed R. 2010. A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nat Commun* **1:** 97. doi:10.1038/ncomms1103

Fei J, Jadaliha M, Harmon TS, Li ITS, Hua B, Hao Q, Holehouse AS, Reyer M, Sun Q, Freier SM, et al. 2017. Quantitative analysis of multilayer organization of proteins and RNA in nuclear speckles at super resolution. *J Cell Sci* **130:** 4180–4192. doi:10.1242/jcs.206854

Femino AM, Fay FS, Fogarty K, Singer RH. 1998. Visualization of single RNA transcripts *in situ*. *Science* **280:** 585–590. doi:10.1126/science.280.5363.585

Frankiw L, Majumdar D, Burns C, Vlach L, Moradian A, Sweredoski MJ, Baltimore D. 2019. BUD13 promotes a type I interferon response by countering intron retention in Irf7. *Mol Cell* **73:** 803–814.e6. doi:10.1016/j.molcel.2018.11.038

Galganski L, Urbanek MO, Krzyzosiak WJ. 2017. Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Res* **45:** 10350–10368. doi:10.1093/nar/gkx759

Gaspar I, Wippich F, Ephrussi A. 2017. Enzymatic production of single-molecule FISH and RNA capture probes. *RNA* **23:** 1582–1591. doi:10.1261/rna.061184.117

Girard C, Will CL, Peng J, Makarov EM, Kastner B, Lemm I, Urlaub H, Hartmuth K, Lührmann R. 2012. Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat Commun* **3:** 994. doi:10.1038/ncomms1998

Grabski DF, Broseus L, Kumari B, Rekosh D, Hammarskjold M, Ritchie W. 2021. Intron retention and its impact on gene expression and protein diversity: a review and a practical guide. *Wiley Interdiscip Rev RNA* **12:** e1631. doi:10.1002/wrna.1631

Hayashi R, Handler D, Ish-Horowicz D, Brennecke J. 2014. The exon junction complex is required for definition and excision of neighboring introns in *Drosophila*. *Genes Dev* **28:** 1772–1785. doi:10.1101/gad.245738.114

Jahn K, Tørring T, Voigt NV, Sørensen RS, Bank Kodal AL, Andersen ES, Gothelf KV, Kjems J. 2011. Functional patterning of DNA origami by parallel enzymatic modification. *Bioconjugate Chem* **22:** 819–823. doi:10.1021/bc2000098

Lanctôt C. 2018. Enzyme-based synthesis of single molecule RNA FISH probes. *Mol Biol* doi:10.1101/248245

Mauger O, Lemoine F, Scheiffele P. 2016. Targeted intron retention and excision for rapid gene regulation in response to neuronal activity. *Neuron* **92:** 1266–1278. doi:10.1016/j.neuron.2016.11.032

McHugh CA, Chen C-K, Chow A, Surka CF, Tran C, McDonel P, Pandya-Jones A, Blanco M, Burghard C, Moradian A, et al. 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521:** 232–236. doi:10.1038/nature14443

Mishra K, Kanduri C. 2019. Understanding long noncoding RNA and chromatin interactions: what we know so far. *ncRNA* **5:** 54. doi:10.3390/ncrna5040054

Monteuuis G, Wong JJL, Bailey CG, Schmitz U, Rasko JEJ. 2019. The changing paradigm of intron retention: regulation, ramifications and recipes. *Nucleic Acids Res* **47:** 11497–11513. doi:10.1093/nar/gkz301

Motea EA, Berdis AJ. 2010. Terminal deoxynucleotidyl transferase: the story of a misguided DNA polymerase. *Biochim Biophys Acta* **1804:** 1151–1166. doi:10.1016/j.bbapap.2009.06.030

Mueller F, Senecal A, Tantale K, Marie-Nelly H, Ly N, Collin O, Basyuk E, Bertrand E, Darzacq X, Zimmer C. 2013. FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat Methods* **10:** 277–278. doi:10.1038/nmeth.2406

Pandya-Jones A, Bhatt DM, Lin C-H, Tong A-J, Smale ST, Black DL. 2013. Splicing kinetics and transcript release from the chromatin compartment limit the rate of Lipid A-induced gene expression. *RNA* **19:** 811–827. doi:10.1261/rna.039081.113

Pandya-Jones A, Markaki Y, Serizay J, Chitiashvili T, Mancia Leon WR, Damianov A, Chronis C, Papp B, Chen C-K, McKee R, et al. 2020. A protein assembly mediates Xist localization and gene silencing. *Nature* **587:** 145–151. doi:10.1038/s41586-020-2703-0

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. 2006. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol* **4:** e309. doi:10.1371/journal.pbio.0040309

Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. 2008. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5:** 877–879. doi:10.1038/nmeth.1253

Rowland TJ, Dumbović G, Hass EP, Rinn JL, Cech TR. 2019. Single-cell imaging reveals unexpected heterogeneity of TERT expression across cancer cell lines. *Cancer Biol* doi:10.1101/618769

Shah S, Lubeck E, Schwarzkopf M, He T, Greenbaum A, Sohn CH, Lignell A, Choi HMT, Gradinaru V, Pierce NA, et al. 2016. Single-molecule RNA detection at depth via hybridization chain reaction and tissue hydrogel embedding and clearing. *Development* **143:** 2862–2867. doi:10.1242/dev.138560

Shah S, Takei Y, Zhou W, Lubeck E, Yun J, Eng C-HL, Koulena N, Cronin C, Karp C, Liaw EJ, et al. 2018. Dynamics and spatial genomics of the nascent transcriptome by intron seqFISH. *Cell* **174:** 363–376.e16. doi:10.1016/j.cell.2018.05.035

Tammer L, Hameiri O, Keydar I, Roy VR, Ashkenazy-Titelman A, Custódio N, Sason I, Shayevitch R, Rodríguez-Vaello V, Rino J, et al. 2022. Gene architecture directs splicing outcome in separate nuclear spatial regions. *Mol Cell* **82:** 1021–1034.e8. doi:10.1016/j.molcel.2022.02.001

Tsanov N, Samacoits A, Chouaib R, Traboulsi A-M, Gostan T, Weber C, Zimmer C, Zibara K, Walter T, Peter M, et al. 2016. smiFISH and FISH-quant: a flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Res* **44:** e165. doi:10.1093/nar/gkw784

Vargas DY, Shah K, Batish M, Levandoski M, Sinha S, Marras SAE, Schedl P, Tyagi S. 2011. Single-molecule imaging of transcriptionally coupled and uncoupled splicing. *Cell* **147:** 1054–1065. doi:10.1016/j.cell.2011.10.024

Wang S. 2019. Single molecule RNA FISH (smFISH) in whole-mount mouse embryonic organs. *Curr Protoc Cell Biol* **83:** e79. doi:10.1002/cpcb.79

Wang X, Liu C, Zhang S, Yan H, Zhang L, Jiang A, Liu Y, Feng Y, Li D, Guo Y, et al. 2021. N6-methyladenosine modification of MALAT1 promotes metastasis via reshaping nuclear speckles. *Dev Cell* **56:** 702–715.e8. doi:10.1016/j.devcel.2021.01.015

Winz M-L, Linder EC, André T, Becker J, Jäschke A. 2015. Nucleotidyl transferase assisted DNA labeling with different click chemistries. *Nucleic Acids Res* **43:** e110. doi:10.1093/nar/gkv544

Wong JJ-L, Ritchie W, Ebner OA, Selbach M, Wong JWH, Huang Y, Gao D, Pinello N, Gonzalez M, Baidya K, et al. 2013. Orchestrated intron retention regulates normal granulocyte differentiation. *Cell* **154:** 583–595. doi:10.1016/j.cell.2013.06.052

Xie F, Timme KA, Wood JR. 2018. Using single molecule mRNA fluorescent in situ hybridization (RNA-FISH) to quantify mRNAs in individual murine oocytes and embryos. *Sci Rep* **8:** 7930. doi:10.1038/s41598-018-26345-0

Yeom K-H, Pan Z, Lin C-H, Lim HY, Xiao W, Xing Y, Black DL. 2021. Tracking pre-mRNA maturation across subcellular compartments identifies developmental gene regulation through intron retention and nuclear anchoring. *Genome Res* **31:** 1106–1119. doi:10.1101/gr.273904.120

Young AP, Jackson DJ, Wyeth RC. 2020. A technical review and guide to RNA fluorescence in situ hybridization. *PeerJ* **8:** e8806. doi:10.7717/peerj.8806

See the following page for ***Meet the First Author***

MEET THE FIRST AUTHOR

Wen Xiao

***Meet the First Author(s)* is an editorial feature within RNA, in which the first author(s) of research-based papers in each issue have the opportunity to introduce themselves and their work to readers of RNA and the RNA research community. Wen Xiao is the first author of this paper, "Improved enzymatic labeling of fluorescent in situ hybridization probes applied to the visualization of retained introns in cells." Wen is a postdoc in Doug Black's laboratory at the University of California, Los Angeles (UCLA). Wen's research focuses on investigating the role of alternative splicing, especially retained introns, and lncRNAs in neuronal development.**

What are the major results described in your paper and how do they impact this branch of the field?

Our method is a modification of a paper published in *RNA* several years ago by Anne Ephrussi's laboratory. With this new method, the labeling is more consistent and can use a greater variety of fluorophores. We describe this method and show that it gives equivalent results to others. We then apply it to examine the location of retained intron transcripts in cells. We find that transcripts exhibiting very different fractionation behavior and different processing of their retained introns present similar patterns of localization for their unspliced RNA. We believe that our method will be valuable for a wide range of groups performing fluorescence in situ hybridization assays.

What led you to study RNA or this aspect of RNA science?

I was fascinated by RNA biology when I was a graduate student in Dr. Yungui Yang's laboratory at the Beijing Institute of Genomics, Chinese Academy of Sciences. During my PhD, I focused on the study of one of the most prevalent modifications on mRNA, m^6A, and found that this modification can regulate the alternative splicing of pre-mRNA by recruiting splicing factors. To pursue an academic career, I joined Dr. Black's laboratory as a postdoc at UCLA to study the mechanism of splicing regulation and the role of lncRNA in brain development.

During the course of these experiments, were there any surprising results or particular difficulties that altered your thinking and subsequent focus?

I encountered a bottleneck when I initially tried to optimize the FISH probe labeling protocol. My mentor Doug Black advised me to change the order of labeling and try more fluorescent dyes. This suggestion worked so well that we can now obtain high labeling efficiency for our FISH probes, which could be used for the detection of pre-mRNAs, lncRNAs, and even retained introns.

If you were able to give one piece of advice to your younger self, what would that be?

Science is hard but intriguing. If I had a chance to give some advice to my younger self, I would say, "Be patient and resilient, be confident with your assumptions, don't be too aggressive and verify it carefully and progressively."

What are your subsequent near- or long-term career plans?

My long-term career plan is to find a faculty position and set up my own laboratory at a research institute/university or find an industrial job in a company.